VISAGE: Visualization for Integrated Satellite, Airborne and Ground-based data Exploration

AIST-16-0094

NASA

ESTO
Earth Science Technology Office

itsc
INFORMATION TECHNOLOGY AND SYSTEMS CENTER

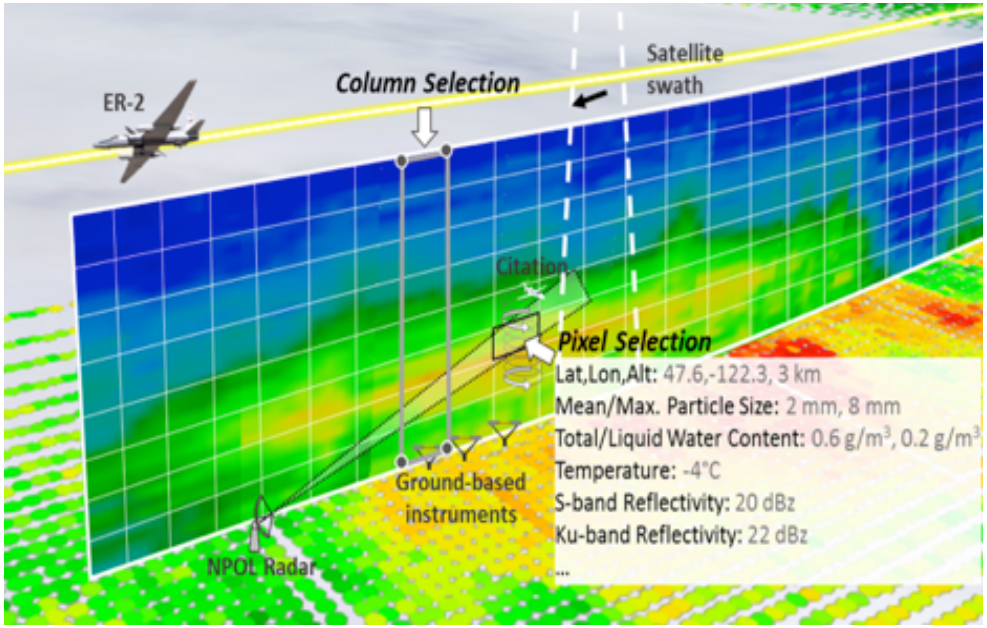
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

VISAGE Team

PI:	*Helen Conover, UAH*
Co-Is:	*Todd Berendes, Aaron Naeger, UAH*
	Manil Maskey, Patrick Gatlin, MSFC
	Stephanie Wingo, USRA/NPP
Team:	*Ajinkya Kulkarni, Charles Collins, Abdelhak Marouane, Lihua Wang, UAH*
Students:	*Shivangi Gupta, Khomsun Singhirunnusorn, UAH*
Collaborators:	*Walter Petersen, MSFC*
	David Wolff, GSFC/WFF

Introduction

- VISAGE is a data exploration tool that will facilitate more efficient Earth Science investigations by providing visualization and analytic capabilities for diverse coincident datasets, with a focus on airborne field campaigns
  - 2D and 3D visualization
  - Data interrogation using map interface
  - Basic analytics (e.g. histograms)



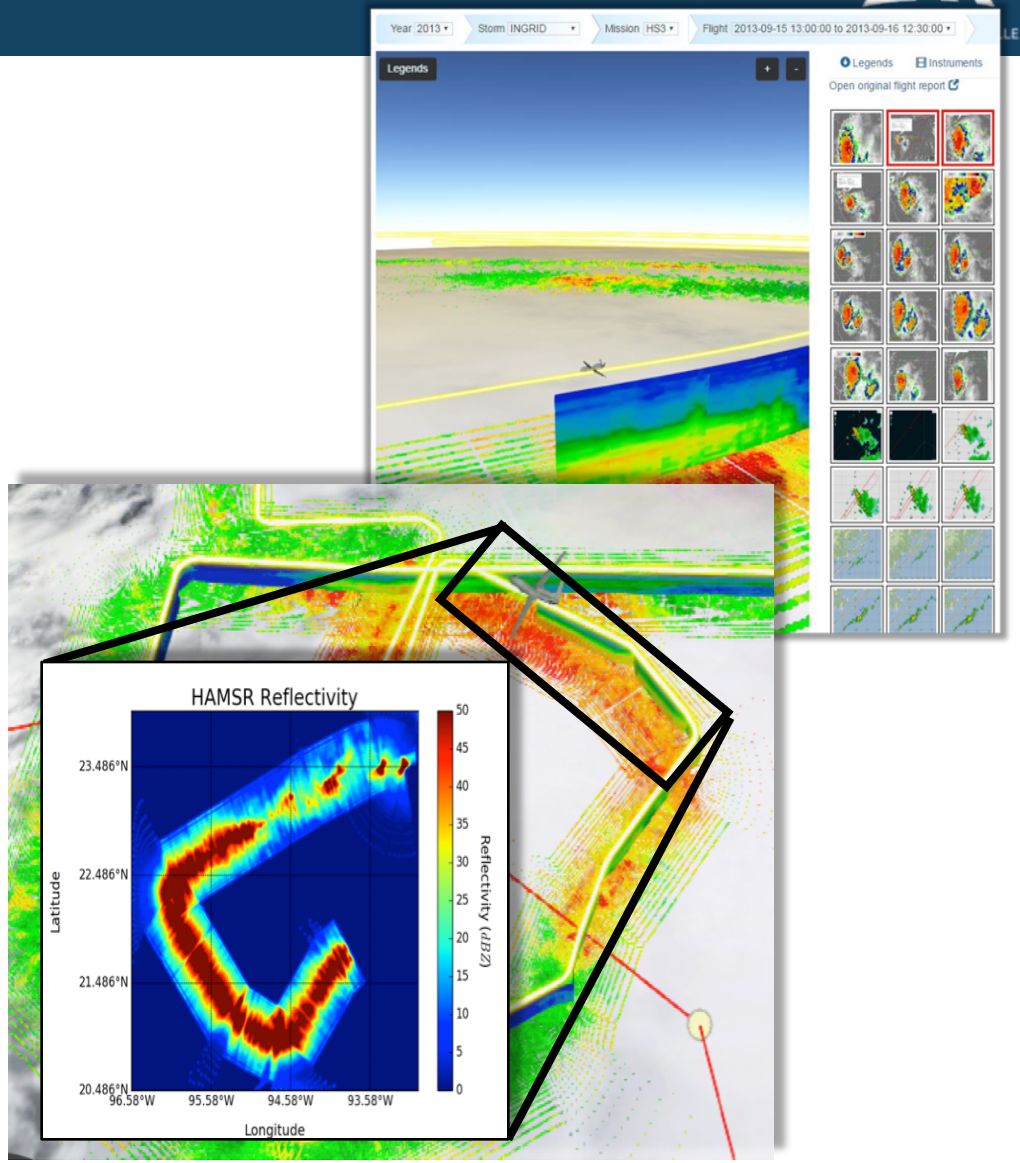
VISAGE concept to visualize and interrogate diverse, fused field campaign datasets

- AIST-funded research project
- Low Technology Readiness Level (TRL) – not ready for production

- *Long term vision – a robust multi-sensor, multi-format integration system suitable for a wide array of applications*

Technology Basis: Field Campaign Explorer

- A data visualization tool for exploring a variety of field campaign data
- Developed around the Hurricane and Severe Storm Sentinel (HS3) Earth Venture mission
- Allows users to
  - Explore missions reports
  - Re-enact mission flights by visualizing data within an interactive environment
  - Subset and preview data on the fly
- Web-based, no software installation required



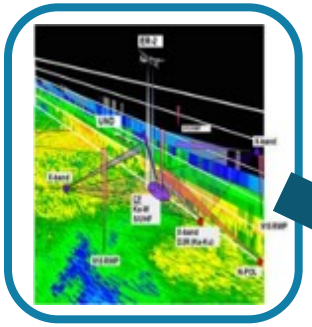
Science Value

- Proof-of-concept to be centered around the ***GPM Ground Validation program***
  - Intensive, coincident observations of atmospheric phenomena
  - Data from a wide variety of ground-based, airborne and satellite instruments
  - Diversity in spatial and temporal scales, variables, formats, etc., makes these data difficult to use together
- VISAGE can bring together these diverse measurements into a common framework to provide
  - Interactive user interface for visualization and analytics
  - Cloud-based VISAGE repository containing data specific to the selected use cases
  - Support for target user community – NASA Precipitation Measurement Mission Science Team
- Focus on "golden cases" where most ground instruments were in operation and multiple research aircraft sampled a significant weather event, ideally while the GPM Core Observatory passed overhead

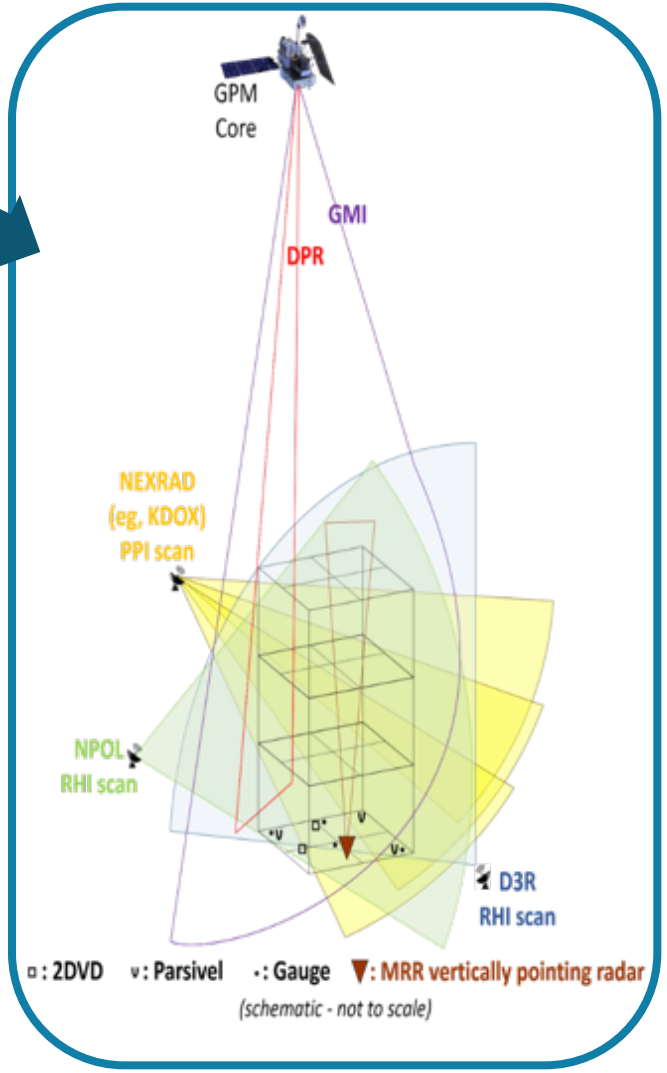
Targeted Data Products: SIMBA

System for Integrating Multi-platform data to Build the Atmospheric column

- Higher-level data product, fusing **GPM satellite and ground-based observations** into a gridded atmospheric column data file
- Interpolates and/or resamples observations from various scales to set data into a common, **user-specified 3D grid**
- Encodes observations from diverse data formats into unified **netCDF file**
- **Attributes** preserve key operation parameters for each sensor
  - Location, operation mode, timestamps, algorithms, product versions, etc.



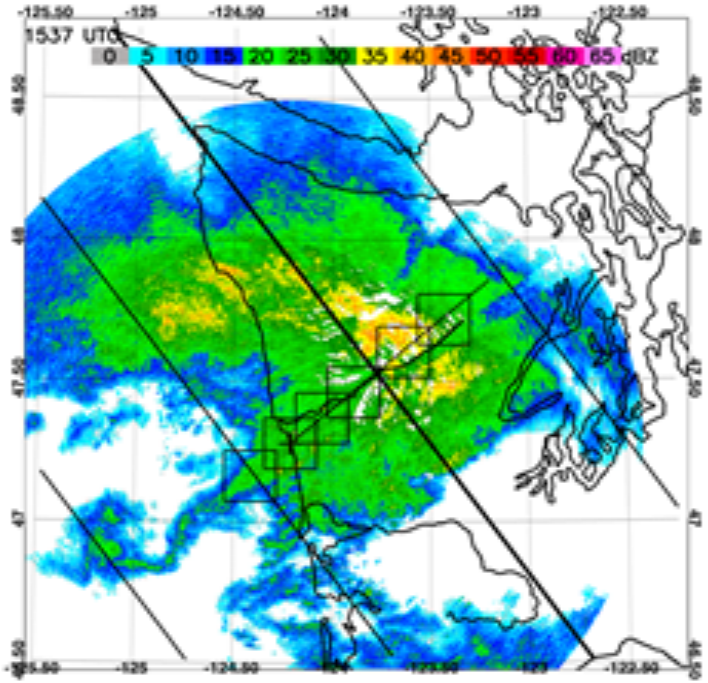
SIMBA fuses targeted satellite- and ground-based observations collected in various formats & coordinate systems to a single, used-defined 3D column grid



Wingo et al. – *J. Atmos. Ocean. Technol.* 2018
doi:10.1175/JTECH-D-17-0187.1

Initial Focus Use Case: 3 Dec 2015 from the OLYMPEX Field Campaign

Complex baroclinic system with orographic enhancement; excellent sampling coordination with simultaneous satellite, airborne, & ground-based observations



3 Dec NPOL reflectivity at 1537, showing GPM DPR and GMI overpass swaths (bold line is nadir track) and select SIMBA column locations

GHRC Data

Data Available for this use case:

- Most ground instruments, including:
  - Radars: **NPOL**, KLGX, D3R, DOW
  - Disdrometers, **gauges**, profilers, soundings
- **ER-2**, DC-8, and Citation aircraft:
  - **AMPR**, CRS, APR-3, CoSMIR, microphysics probes, etc
- **GPM VN (DPR and ground radar match-ups)**
- **GPM** overpass at 15:22 UTC
  - GMI swath images from GIBS
- Select **SIMBA** columns
- Select **WRF** model subsets

*Data in **bold** in VISAGE repository*

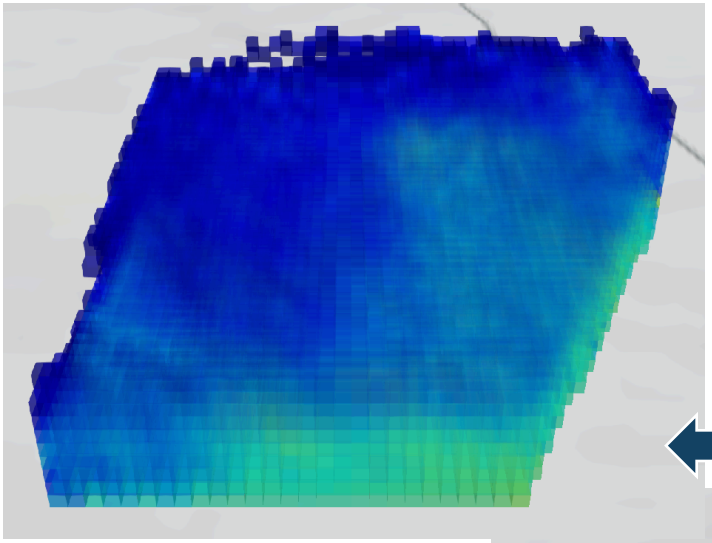
Science question: How well does the WRF model (P3 microphysics) depict the precipitation enhancement from the ocean to the land and across the Olympic mountains?

Technical Challenges / Research Areas

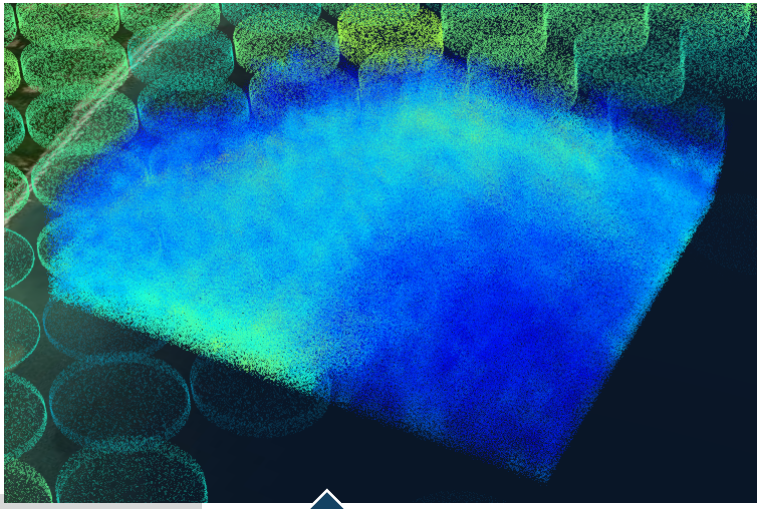
- **3D data visualization and exploration of large data volumes on a web-based platform**
  - **Cesium** open source geospatial 3D mapping platform
  - Investigating **multiple 3D rendering** approaches
  - Data interrogation via map user interface
  - Basic analytics across different data sources (e.g., histograms)
- **Analysis ready data repository**
  - Apache **Parquet** data storage format supporting very efficient compression and encoding schemes
- **Serverless Cloud-Native Technologies**
  - Amazon Web Service (AWS) **Athena** stateless query service for searching data stored in S3 buckets
  - AWS **Step Functions** and **Lambdas** to orchestrate and run data processing and rendering code without provisioning or managing servers, automatically scaling resources as needed
  - Aligning with ESDIS's Cloud activities

Different 3D Rendering Approaches

NPOL radar reflectivity, gridded in a SIMBA column. Comparing time to render, browser memory usage, appearance

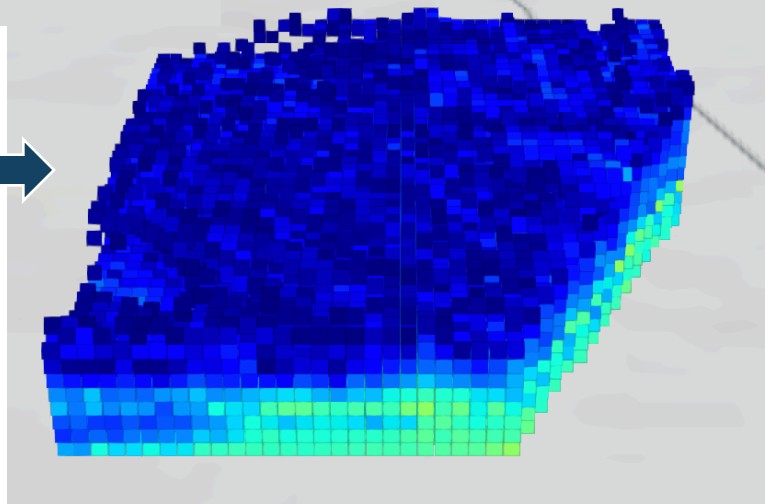


Grid cells, color values, transparency defined using Cesium's native "entities" with CZML



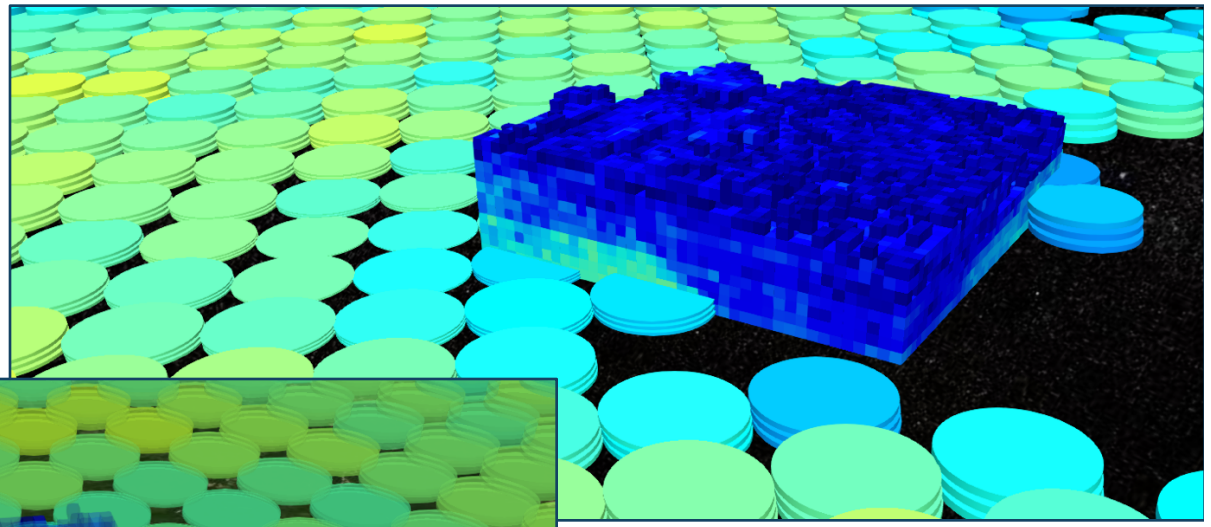
Point cloud, with transparency weighted by dbz values (higher values are less transparent to accentuate "hot spots")

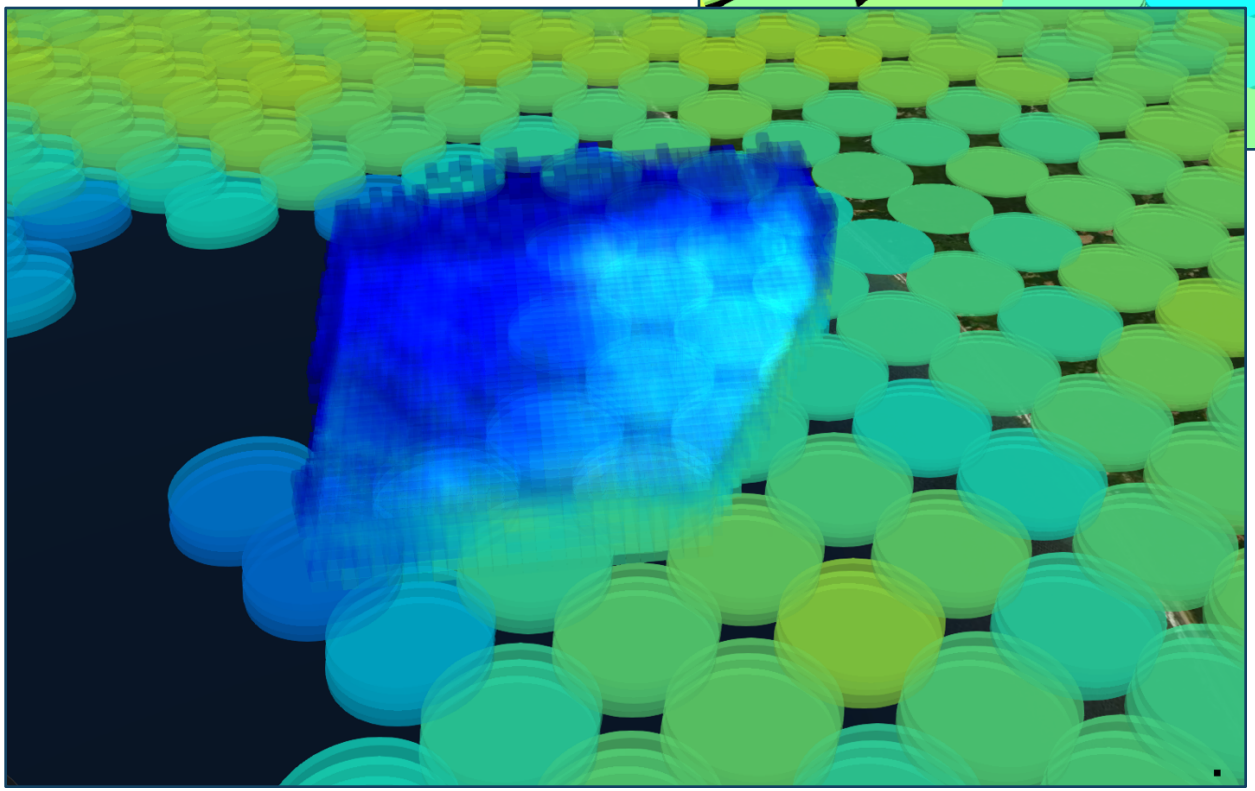
Grid cells rendered as 2D "billboards" with CZML



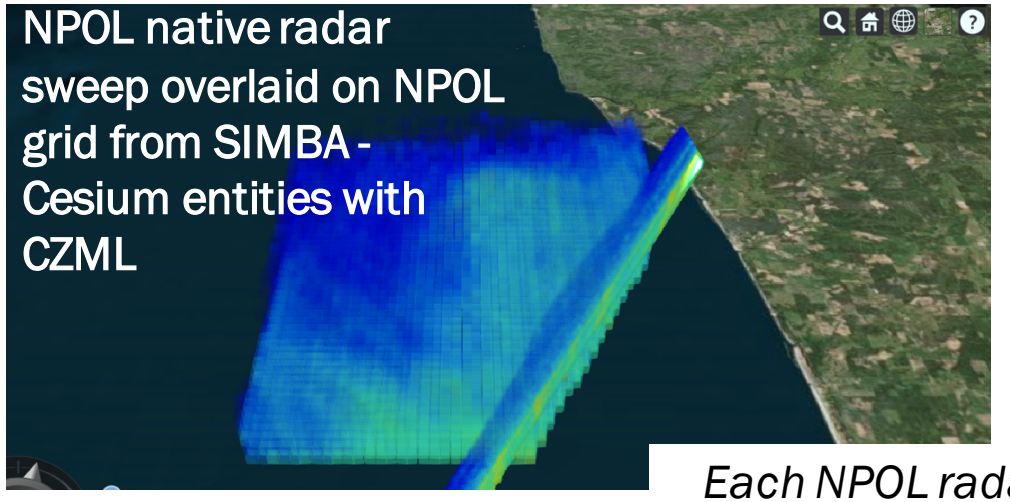
Solid vs. Transparent Objects

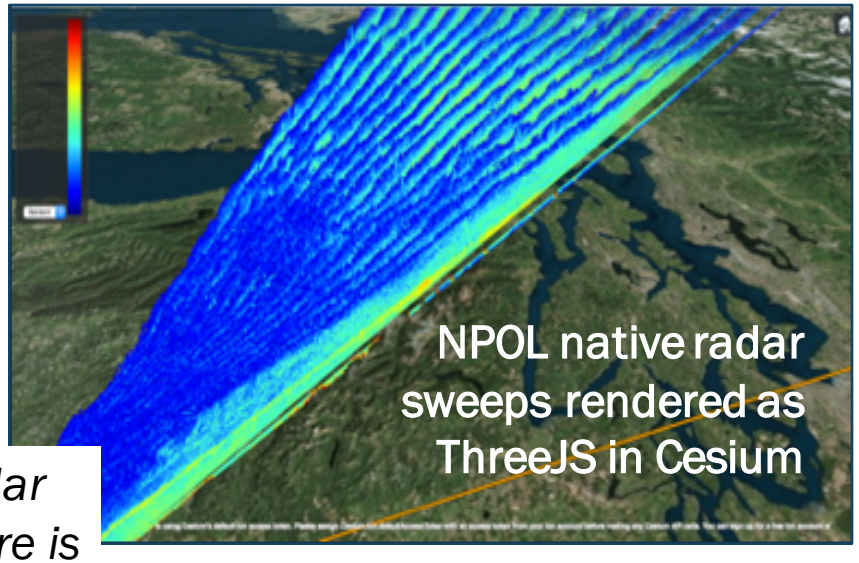
Validation Network (cylinders) overlaid with NPOL grid from SIMBA column (cubes) – Cesium entities with CZML



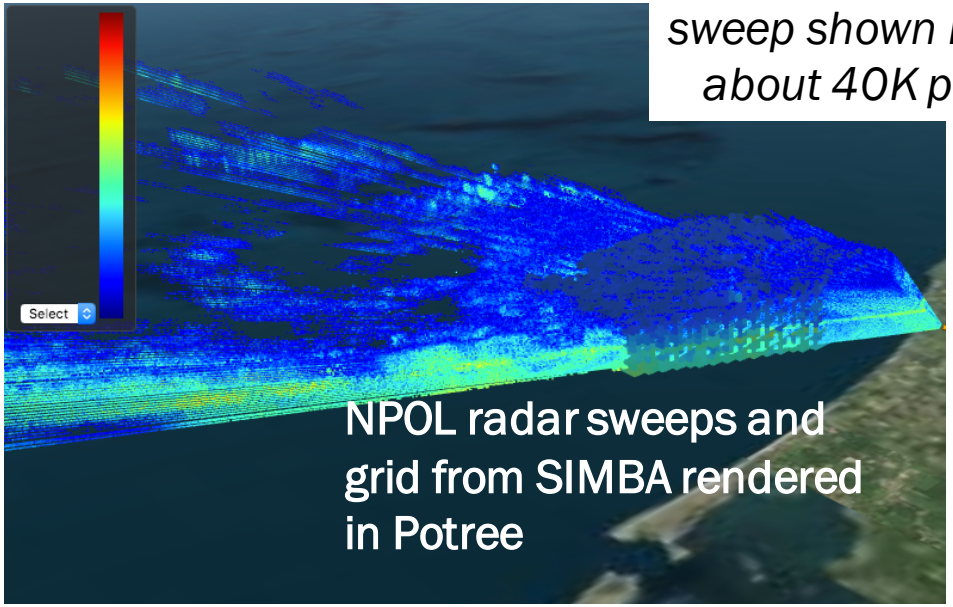


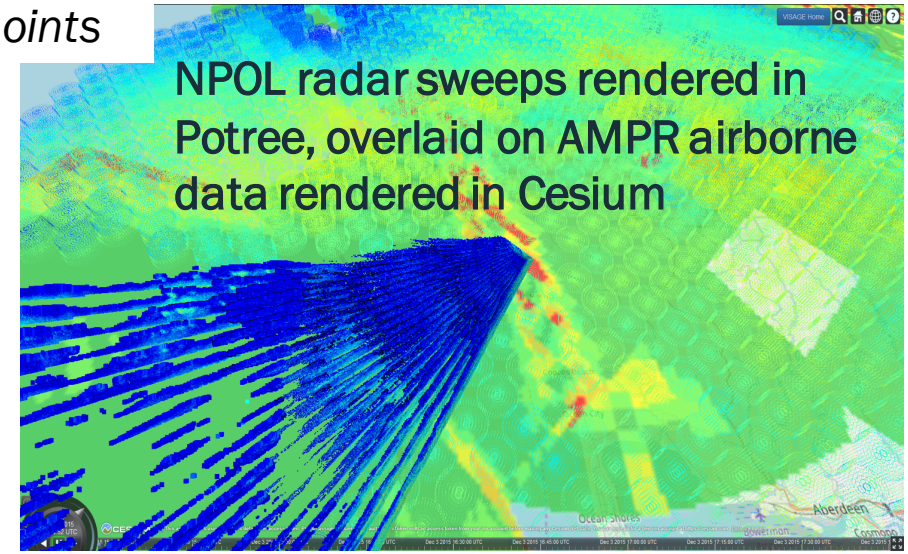
Experiments with Radar Reflectivity





Each NPOL radar sweep shown here is about 40K points





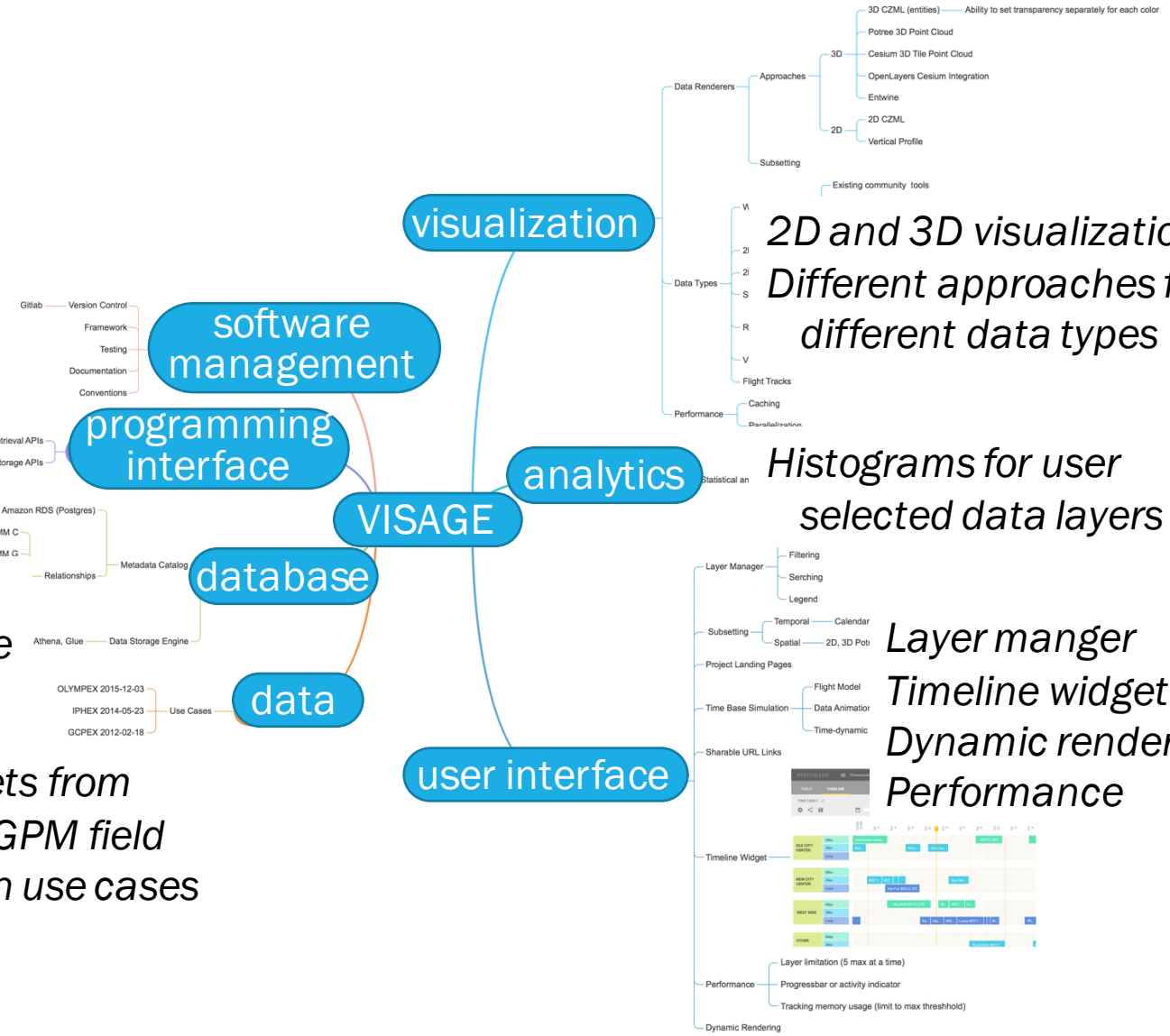
User Scenarios

- Nominal Work Flow
  - Select from pre-defined golden cases
  - U/I will zoom to region and time of interest; display default layers for use case
  - Use layer manager to select among additional data layers, grouped by data variable
  - Visually compare observations from different instruments or between model and observation
  - Mouse over data points of intense activity to see data values
- Alternate data explorations:
  - Look for when there are the most observations at the same time
  - Start with SIMBA column and explore source observations
  - Start with WRF and look at related observations
  - Start with rain gauge to look for intense precipitation event
  - Start with flight track and explore observations from airborne instruments
  - Compare WRF and SIMBA data cubes

VISAGE demo

BACKUP SLIDES

VISAGE Components



Version control
Testing
Documentation

Data storage & retrieval

Metadata catalog
Data storage engine

Key data sets from selected GPM field campaign use cases

2D and 3D visualizations
Different approaches for different data types

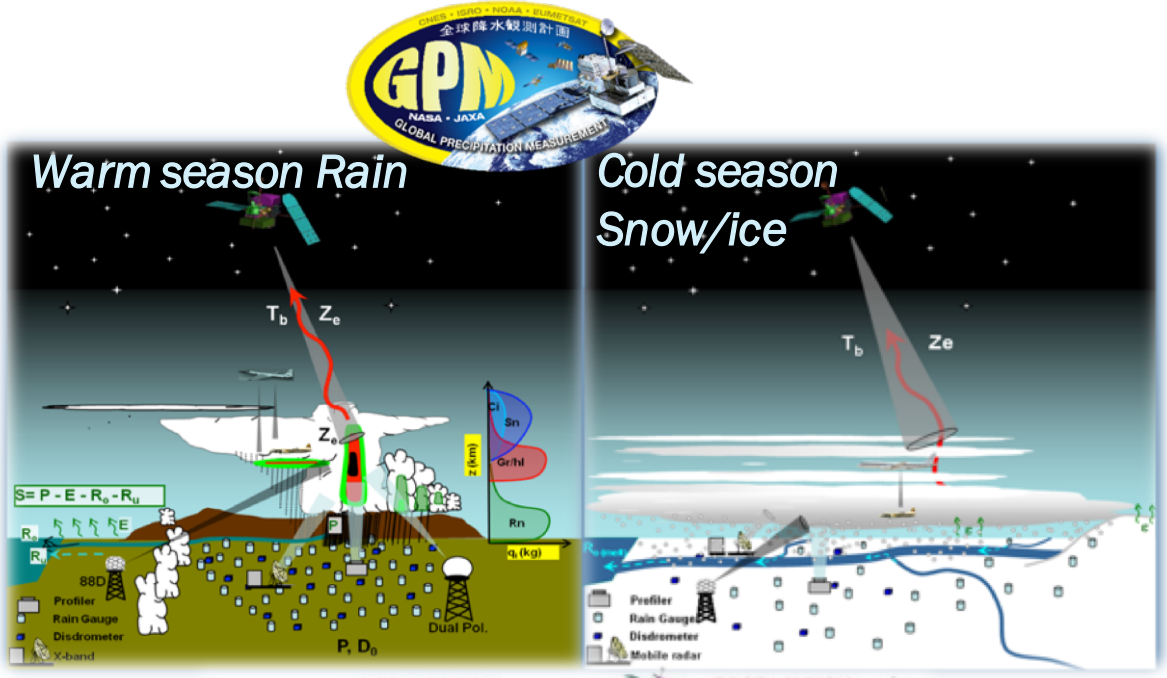
Histograms for user selected data layers

Layer manger
Timeline widget
Dynamic rendering
Performance

Targeted Data Products: GPM Ground Validation Archive

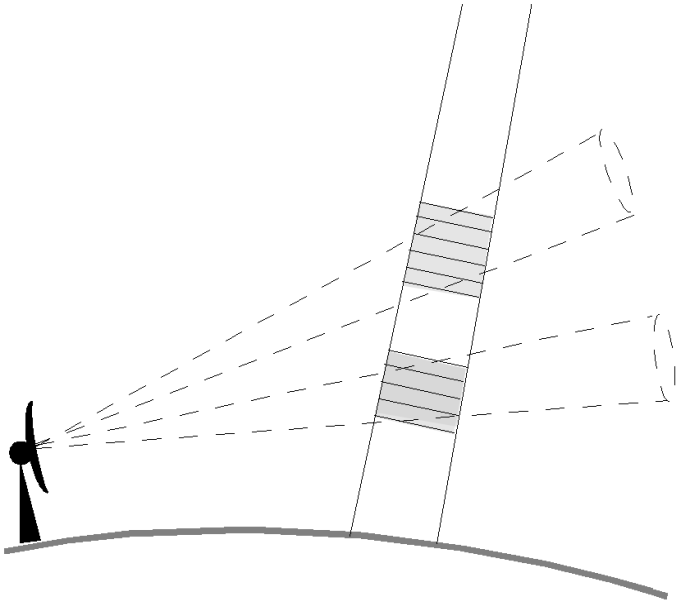
GPM GV data are archived at the GHRC DAAC. This collection includes:

- Series of field campaigns collecting detailed measurements of precipitation and related physical processes in a variety of diverse metrological regimes
- Ground and airborne precipitation datasets supporting validation of satellite-based precipitation retrieval algorithms
- Related extended observations from additional sites



Targeted Data Products: GPM Validation Network (VN)

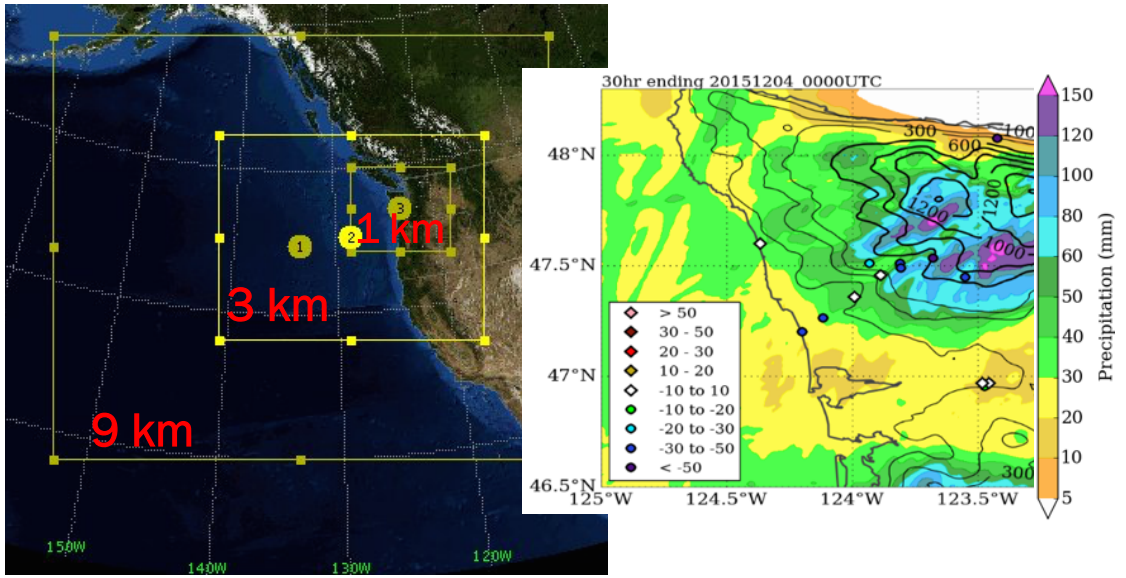
- Compares data from satellite radars (GPM DPR, TRMM PR) and microwave imagers (GMI, TMI and others) to ground-based scanning weather radar observations
- Subsets satellite and ground radar (GR) data for coincident observations of precipitation
- Generates vertical profiles with matching coincident DPR and GR data for precipitation events



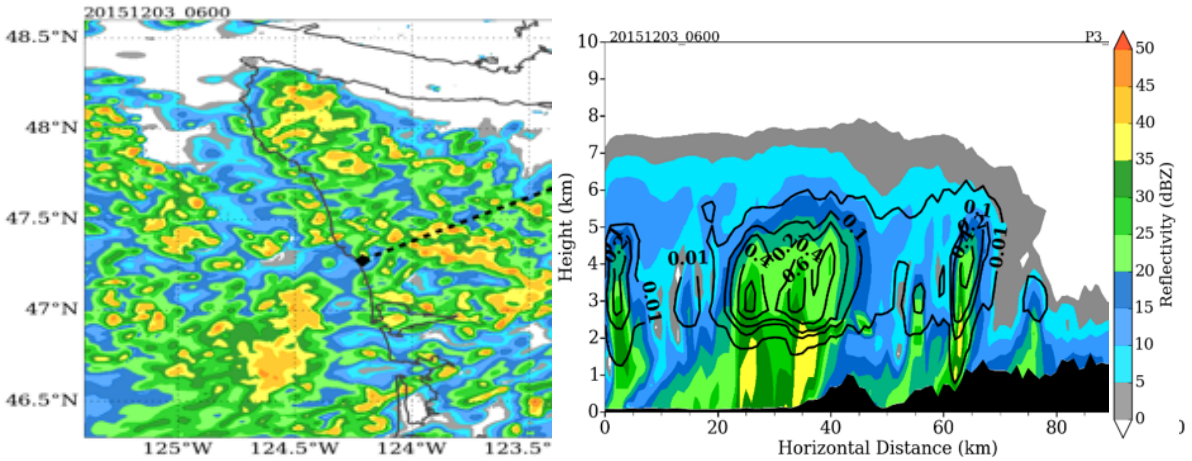
Example of DPR gate averaging at GR sweep intersections. Shaded areas show individual DPR gates intersecting the vertical extent of two GR sweeps (dashed) at different elevation angles. The reflectivity values of the individual DPR gates are averaged over the vertical extent of the GR sweeps, resulting in two matching volumes for the single DPR ray shown in this case.

WRF Model Simulations

- Triple-nested 9, 3, and 1 km grid setup with high-resolution innermost nest over OLYMPEX field site
- Cloud microphysical schemes in WRF model can lead to large uncertainties in the precipitation forecasts
- Field campaign data can help intensively validate cloud microphysical schemes, but collecting/analyzing the large amount of field and model data can be tedious
- VISAGE will help promote more efficient model validation work, which can ultimately help improve precipitation forecasts



(left) WRF model grid setup, (right) model precipitation versus ground-based observations



(left) WRF model reflectivity at lowest level over the Olympic peninsula, (right) WRF model reflectivity (shaded) and rime mass (solid black) along NPOL RHI scan (black dashed)

Science Use Cases from the GPM GV Program

Olympic Mountains Experiment (OLYMPEX) - Rain & snow in extreme coastal & topographic gradients (NW Washington, Nov 2015 – Feb 2016)

- *Use Case: complex baroclinic system with orographic enhancement on 3 Dec 2015; excellent sampling coordination with simultaneous satellite, airborne, & ground-based*

Integrated Precipitation and Hydrology Experiment (IPHEX) - Warm season precipitation and hydrologic processes in complex terrain (W North Carolina, April – June 2014).

- *Use Case: warm-season convective storm with severe hail on 23-24 May 2014; observations from ground-based radars, two aircraft, and GPM Core satellite with very good GMI and DPR coverage*

GPM Cold season Precipitation Experiment (GCPEX) - Microphysical properties of precipitating snow (Ontario, Canada, Jan – Feb 2012)

- *Use Case: Microphysical observation & simulation of the entire life cycle of a significant precipitation band along a warm front; multiple airborne and ground observations on 18 Feb 2012 (before GPM Core satellite launch in Feb 2014)*

Evolution of 3D Tile Software

Dec 2017 → Mar 2018 → June 2018

py3dtiles – public domain Python module to create point cloud 3dTiles

https://github.com/Oslandia/py3dtiles

CesiumTiles – locally developed Java function implements additional features of the point cloud 3dTile

VisageTiles – lambda function to generate 3D tiles on-the-fly and render them in Cesium. Multiple VisageTiles lambdas can be invoked in parallel

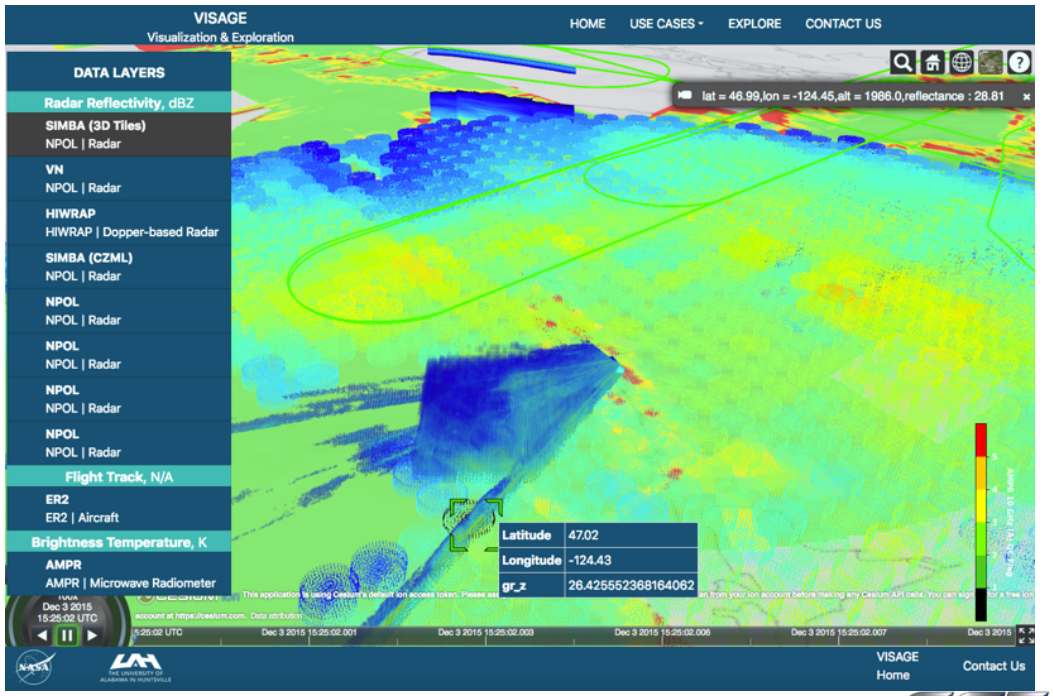
	py3dtiles	CesiumTiles	VisageTiles
# points	3.9 million	3.9 million	3.9 million
# tiles	2500	1	4
Time to create	13+ min	11 sec	6 sec total elapsed time
Time to display	30+ min	<1 sec	

Note: time to create 3D tiles is related to number of points, while time to display depends more on number of tiles

User Interface – Web Application

- VISAGE is a serverless web application hosted on an AWS S3 bucket.
- Amazon Gateway API is used to communicate between Cesium user interface and service endpoints
  - Developed using the Angular 6 framework
  - Retrieve metadata and visualizations for display
  - Lambda and Step Functions perform database queries and tileset generation
  - Public API will allow other users to search and download metadata and data

- Cesium JavaScript library is used for data and map visualization
- A layer manager allows users to select and manipulate datasets loaded in Cesium
- Mouse-over for data interrogation.



Cloud Native Technologies: AWS (mostly) Serverless Platform

S3 – Simple Storage Service - within S3, data objects are stored in buckets

Lambda – provides capability to run code without provisioning or managing servers, with automatic scaling

Step Functions – used to coordinate components and step through the functions of an application, e.g., to orchestrate Lambda functions

Athena – serverless, interactive query service analyze data in S3 using standard SQL

Amazon API Gateway – service to create, publish, maintain, monitor, and secure APIs

AWS CloudFormation – tools to describe and provision all the infrastructure resources in the cloud environment

AWS RDS – Relational Database Service supporting different database engines

Apache Parquet – columnar data storage format supporting very efficient compression and encoding schemes